

# A Multicomputer Communications System

J. W. Layland and W. Lushbaugh  
Communications Systems Research Section

*A general description is given of the requirements for, and one proposal for the provision of, the multicomputer communications facility needed in a multiple minicomputer system, such as in the anticipated future tracking station computer network of the Deep Space Network. The main features are: (1) a basically high-speed point-to-point link whose rate adapts without data loss to the capabilities of the computers with which it interfaces; and (2) a very high bandwidth transmission control unit (TCU) which provides a functional path from each computer to every other computer, while requiring only one physical link between each computer and the TCU.*

## I. Introduction

One of the most likely configurations for the computational complement of the tracking stations of the future is that of a number of small computers, each one performing a dedicated function, interfaced with some specific subset of the tracking station equipment; each one would, in turn, be interfaced with the other computers. A similar multiple-computer mini-network is under consideration for the DSN Control System, with each of the included small computers performing a dedicated function. Variants on this configuration have been used successfully at other facilities in a variety of process control and data acquisition systems.

In multicomputer systems such as these, the facility provided for communication between the distinct computers and the usage made of such a facility play a crucial part in their success or failure. If the communications interface between the computers is awkward or if it is too heavily used, the computers may spend significant portions of their capacity in meaninglessly manipulating intercomputer messages, much as a too-large

committee of people can spend more of its efforts in discussing who does what than in progressing toward its goal.

Intercomputer communications are expected to serve two purposes in the multicomputer systems of the DSN. The first is the transient operation of loading programs at the start of operation, or change of mode, plus the loading of diagnostic programs when failures are detected. The maximum conceivable data rate required between computers in this mode is the average data rate from the tapes or disks on which the software semipermanently resides. With current technology, this rate is on the order of  $10^5$  8-bit bytes/s.

The minimum acceptable data rate for this mode is the rate which will allow the system computers to be reloaded in a time-span which meets the up-time/recovery requirements on the system. For example, a system with 30 small computers with 32K of 16-bit words each could be fully loaded in slightly less than 1 min from a single data source at  $5 \times 10^4$  bytes/s.

The second and primary purpose of the intercomputer communications is the relaying of the information needed by the system in the performance of its assigned real-time tasks. The sizing of this capability in the general sense is very difficult, because it depends strongly upon how the functions to be performed are partitioned between the various computers. Nevertheless, an intercomputer data rate which is significantly in excess of the feed-through data rate of the total system probably indicates that the functions were partitioned wrong. For the immediate future, the feed-through data rate will be less than  $5 \times 10^4$  bits/s, the rate of the wide-band data lines (WBDLs). Consequently, a reasonable objective for the intercomputer communications capability is  $5 \times 10^4$  bytes/s. This objective is readily achievable over the short distances expected within a multicomputer system.

## II. The Individual Communications Link

The individual links of the system are used to connect the computers of the mini-network to a centralized switching system, or, alternatively, can be used to connect computer pairs together. This section will describe the characteristics of an individual link, with discussion of multicomputer connections deferred until later.

A typical individual link is diagrammed as Fig. 1. The link functions in a half-duplex mode; that is, it can transmit data in either direction, but only one way at a time. Signals on the two coaxial cables are balanced and transformer-isolated to minimize effects of ground level imbalances between the two computers. The two cables each carry signals in one direction only: one going from computer A to computer B, and the other in the reverse direction. The two line interface units (LIUs) in the figure are identical and contain the logic and analog circuitry needed to format the computer's data byte-by-byte for the lines, perform parity checking, etc. The computer interface units (CIUs) are similar in function but specialized in design to the type of computer to which they are attached.

Each of the four modules occupies between 45 and 60 dual in-line (DIP)-packaged integrated circuit elements. Two CIUs have been designed to date: one for the SDS-910 and one for the PDP-11. Both are interfaced in a Central Processing Unit (CPU) controlled word-by-word transfer mode. Interface to the computers via a direct memory access (DMA) channel could be done at the expense of additional circuitry and the associated design and test effort. This would decrease the load imposed upon the CPU by the communications link.

The three signals which are sent on the lines in the active state are shown in Fig. 2. When the line is inactive, there is no signal. The transmission of the carrier signal is under control of the computer to which the LIU is interfaced. From an inactive state, the switching on of the carrier by computer A is a 1-bit message to computer B that some communication is needed, and the switching on of computer B's carrier acknowledges that message. Symmetrically, the switching off of the carrier is an emphatic "I quit!" that could be used to reinitialize from quasi-catastrophic link operation. These signals used are essentially identical to those of the IBM 2790 loop system (Ref. 1), although the line protocol differs. They admit very straightforward self-synchronization and detection so that no parallel clock lines are needed.

Transmission over a link is, in fact, bit-serial, but it is conceptually on a byte-by-byte basis. The LIU transmits data in a 1-byte packet, together with a parity bit which is checked and stripped by the receiving LIU. Whenever it is able to store another byte in its CIU buffer, the receiving LIU will signal a request for the next byte with a 2-bit control packet. Computer-word buffering is provided in the CIU, and since the LIU transmits by bytes, it is implicitly assumed that the computer word-size is a multiple of 8 bits (one byte). This is valid for the SDS 900-series and Xerox Sigma-series machines, for IBM 360, and for most of the popular 16-bit minicomputers.

The "transmit-byte/receive-control packet" operation sequence provides synchronization between the communication controlling software in both computers that accommodates the link data rate to the capabilities of the computers. Unlike a link built around a self-clocked synchronous modem, no data can be lost, or erroneously inserted, if the computers do not respond within a preset time interval.

This flexibility is obtained at the expense of data rate. The basic signalling frequency of  $10^6$  bits/s is reduced to about  $3 \times 10^5$  bits/s between LIUs over 300 m ( $10^3$  ft) of cable. The rate is still lower when delays in transferring words in and out of the computers is added. What is gained, however, is the ability to set the priority of the intercomputer communication link to that priority which is warranted by the function it performs for the particular subsystem. Since the maintenance of the intercomputer communications link is not the prime job of the subsystem computer, but the link is one of the resources utilized by that computer in performing its prime job, this flexibility is highly desirable.

For example, in a programmed local oscillator (PLO) subsystem, the closed-loop control of the synthesizer and the voltage-controlled oscillator (VCO) can be performed on a high-priority guaranteed service basis, without interference from the intercomputer link which is periodically supplying the PLO subsystem with updated predict information. If the link had a fixed data rate, it might well have to be operated at a higher priority than the PLO equipment in order to prevent rate errors from occurring during the time intervals when the link is active. The impact of this fixed rate constraint on the PLO software could range from annoying to catastrophic, depending upon how difficult the basic PLO system is to implement without the constraint.

### III. Mini-networks of Computers

A myriad of options appears when one considers the interconnection of a number of computers instead of a single pair as discussed above. One obvious answer is to provide a physical link between each pair of computers which has need to communicate. This configuration would entail  $N(N-1)/2$  links for  $N$  computers. It is highly expensive in hardware, but provides the maximum flexibility. Messages directed from one computer to another are directed by addressing the physical link interface hardware attached to the source computer.

We may remove some of the links, connect each computer to only a small number of its neighbors, and yet require that functional data paths exist between computers which are not physically directly connected, provided that message-routing information is contained within the message format itself. The Advanced Research Projects Agency (ARPA) Network (Ref. 2) is of this type. If this elimination of links is carried to its maximum, the result is a star configuration with a message-switching/relaying computer at the center, with each of the subsystem computers linked to it by the points of the star. A network used at IBM manufacturing for process control/data acquisition is of this type (Ref. 3). This star configuration has the disadvantage, relative to the others mentioned, that only one message can be en route at any one time. It also happens to be a functional analog of a single-channel loop communications system where messages enter the loop on a demand basis. The important difference between the star and the analogous loop is that control is concentrated in the center of the star while it would be distributed among the (enlarged) LIUs of a loop system.

If the star network were to be implemented as described, the message switching task could almost certainly be performed by a minicomputer which could comfortably handle two to five messages simultaneously, if it had no other tasks to perform. If the number of subsystem computers were large, then the data rate through this central minicomputer would be the constraining feature of the mini-network.

The operations required at star center, however, are simple and repetitious, and, except possibly for the initial set-up of a message channel based on the message addressing information, could be easily designed indeed into a hardwired transmission control unit (TCU). If built with current transistor-transistor logic (TTL) integrated circuits, the envisioned TCU would have an effective bandwidth of well over  $10^7$  bits/s which would be shared on a demand basis among as many messages as could be established. At this time, the TCU has been only functionally designed, with detailed design postponed until after the LIUs have been thoroughly tested and several CIUs have been designed, built, and tested.

### IV. Message Format

Since the message format utilized by a computer network such as the one described here is embedded in its software, it can be a matter of considerable evolution as experience is gained with the system. At this point there seems to be no valid reason to deviate significantly from ASCII<sup>1</sup> standards in current use. The suggested format skeleton is shown in Fig. 3. The control codes are all ASCII characters, as are the addresses. The addressing is by computer ID, with various tasks within that computer being isolated by a channel identifier. Many more flexible techniques have been proposed (Refs. 4 and 5) but this is sufficient for the relatively predictable real-time jobs of the DSN.

One highly desirable departure from a pure ASCII standard is not shown in Fig. 3. This is the inclusion of the provision for 8-bit binary transmission for programs, unmodified Ground Communications Facility (GCF) data blocks, or in fact almost everything not destined for human reading. The binary data are signalled by an "SO" control code appearing in the message body. The escape-to-binary is followed by the record byte count (one record is of length 1-255) and then the binary data. The binary data are error-protected by the parity bit used by the LIUs.

<sup>1</sup>American Standard Code for Information Interchange.

## V. Summary and Open Problem

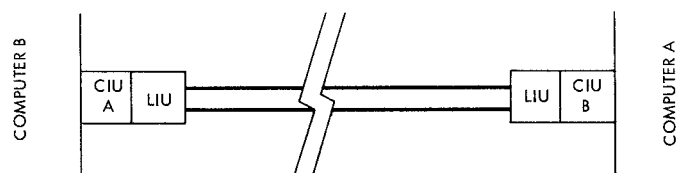
The foregoing has been a general description of the requirements for, and one proposal for the provision of, the multicomputer communications facility needed in a multiple minicomputer system, such as that for the anticipated future tracking station computer network. The main features are: (1) a basically high-speed point-to-point link whose rate adapts without data loss to the capabilities of the computers to which it interfaces; and (2) a very high bandwidth transmission control unit (TCU) which provides a functional path from each com-

puter to every other computer, while requiring only one physical link between each computer and the TCU.

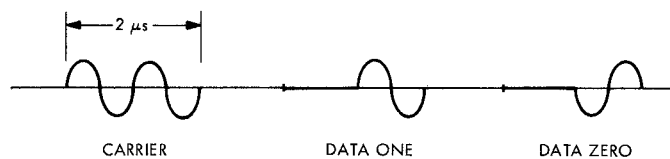
The primary open problems with the proposed configuration center about the capabilities of the TCU; in particular, one must consider the probability of message blocking because the desired destination is already busy. There is a possibility of deadlock if blocked messages are not properly handled. The effects on mini-network behavior of changing the TCU bandwidth, of adding message-storing capability to the TCU, etc., must be assessed.

## References

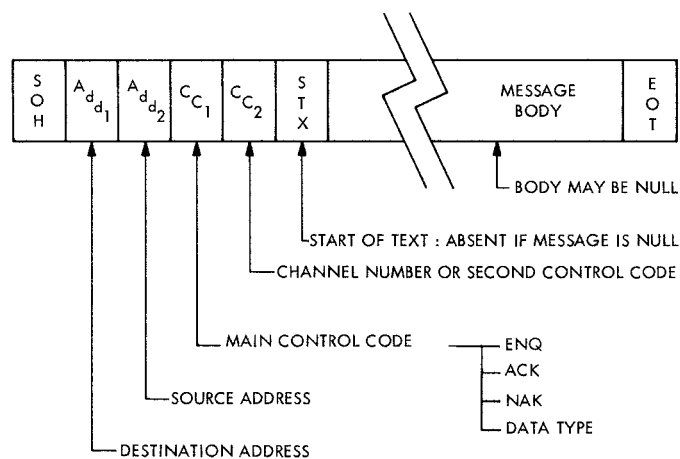
1. Hippert, R. O., "IBM 2790 Digital Transmission Loop," *IBM J. Res. Develop.*, Vol. 14, No. 6, pp. 662-667, Nov. 1970.
2. Roberts, L. G., and Wessler, B. D., "Computer Network Development to Achieve Resource Sharing," *AFIPS Conference Proceedings*, Vol. 36, 1970, pp. 543-549, 1970 Spring Joint Computer Conference.
3. Stuehler, J. E., "An Integrated Manufacturing Process Control System: Implementation in IBM Manufacturing," *IBM J. Res. Develop.*, Vol. 14, No. 6, pp. 605-614, Nov. 1970.
4. Carr, S., Crocker, S. D., and Cerf, V., "HOST-HOST Communication Protocol in the ARPA Network," *AFIPS Conference Proceedings*, Vol. 36, 1970, pp. 589-597, 1970 Spring Joint Computer Conference.



**Fig. 1. Block diagram of single link**



**Fig. 2. Line signals**



**Fig. 3. Proposed message format**